Computational Chemistry Beyond Petascale

Jeff Hammond

Argonne Leadership Computing Facility

10 November 2010

Background

2009-present	—	Argonne Director's Fellowship
Summer 2006	—	NWChem internship
2005-2009	—	DOE-CSGF
2003-2009	—	UChicago Chemistry

THE UNIVERSITY OF
CHICAGO

DOE CSGF

NWCHEM
HIGH-PERFORMANCE COMPUTATIONAL
CHEMISTRY SOFTWARE

Outline

1. Brief overview of computational chemistry
2. Why exascale matters to chemists
3. Four examples of massive parallelism transforming chemical applications

Computational chemistry

Atomistic simulation in chemistry

1. classical molecular dynamics (MD) with empirical potentials
2. ab initio molecular dynamics based upon density-function theory (DFT)
3. quantum chemistry with wavefunctions e.g. coupled-cluster (CC)

Classical molecular dynamics

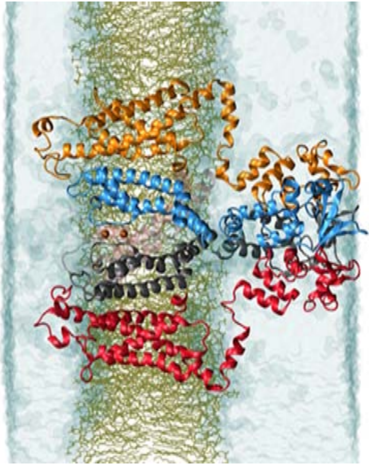


Image courtesy of Benoît Roux via ALCF.

- Solves Newton's equations of motion with empirical terms and classical electrostatics.
- Size: 100K-10M atoms
- Time: 1-10 ns/day
- Scaling: $\sim N_{atoms}$

Data from K. Schulten, et al. "Biomolecular modeling in the era of petascale computing." In D. Bader, ed., *Petascale Computing: Algorithms and Applications*.

Car-Parrinello molecular dynamics

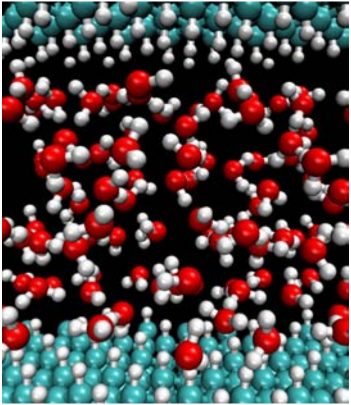
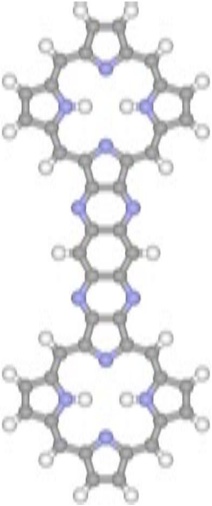


Image courtesy of Giulia Galli via ALCF.

- Forces obtained from solving an approximate single-particle Schrödinger equation; time-propagation via Lagrangian approach.
- Size: 100-1000 atoms
- Time: 0.01-1 ns/day
- Scaling: $\sim N_{el}^{x}$ (x=1-3)

F. Gygi, *IBM J. Res. Dev.* **52**, 137 (2008); E. J. Bylaska et al. *J. Phys.: Conf. Ser.* **180**, 012028 (2009).

Coupled-cluster theory



- Infinite-order solution to many-body Schrödinger equation truncated via clusters.
- Size: 10-100 atoms
- Time: N/A
- Scaling: $\sim N_{bf}^{x}$ (x=4-7)

Image courtesy of Karol Kowalski and Niri Govind.

Chemistry on supercomputers

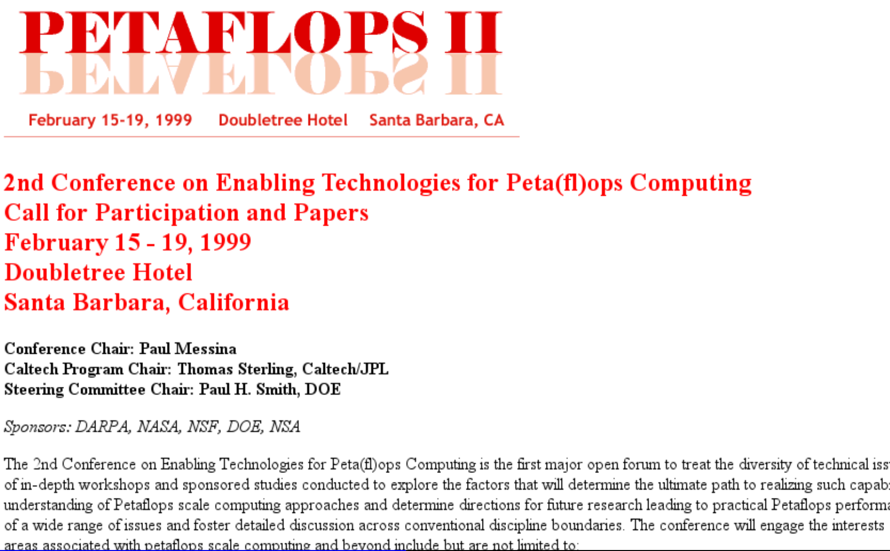
Both classical and ab initio molecular dynamics have essentially reached algorithmic maturity. Most research is fighting Amdahl's law and related concepts (FFT does not scale), e.g. DEShaw has turned classical molecular dynamics into an engineering problem.

Quantum many-body methods are far from algorithmic maturity because they have been constrained to tiny systems so the N-body problem is hidden behind dense linear algebra.

Dense linear algebra is great for Gordon Bell Prizes but terrible for science.

Why is exascale different *for chemists*?

Deja Vu I



PETAFLOPS II

February 15-19, 1999 Doubletree Hotel Santa Barbara, CA

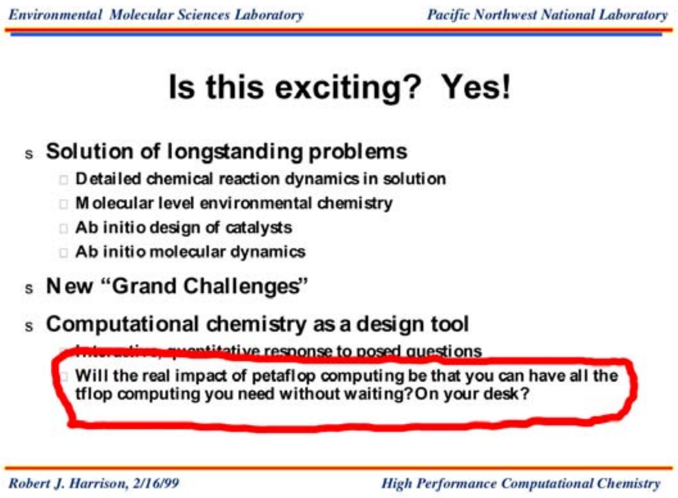
2nd Conference on Enabling Technologies for Peta(fl)ops Computing
Call for Participation and Papers
February 15 - 19, 1999
Doubletree Hotel
Santa Barbara, California

Conference Chair: Paul Messina
Caltech Program Chair: Thomas Sterling, Caltech/JPL
Steering Committee Chair: Paul H. Smith, DOE

Sponsors: DARPA, NASA, NSF, DOE, NSA

The 2nd Conference on Enabling Technologies for Peta(fl)ops Computing is the first major open forum to treat the diversity of technical iss
of in-depth workshops and sponsored studies conducted to explore the factors that will determine the ultimate path to realizing such capab
understanding of Petaflops scale computing approaches and determine directions for future research leading to practical Petaflops performa
of a wide range of issues and foster detailed discussion across conventional discipline boundaries. The conference will engage the interests
areas associated with petaflops scale computing and beyond include but are not limited to:

Deja Vu II



Environmental Molecular Sciences Laboratory *Pacific Northwest National Laboratory*

Is this exciting? Yes!

- Solution of longstanding problems
  - Detailed chemical reaction dynamics in solution
  - Molecular level environmental chemistry
  - Ab initio design of catalysts
  - Ab initio molecular dynamics
- New "Grand Challenges"
- Computational chemistry as a design tool
  - Interactive, quantitative response to posed questions
  - Will the real impact of petaflop computing be that you can have all the tflop computing you need without waiting? On your desk?

Robert J. Harrison, 2/16/99 *High Performance Computational Chemistry*

Why is computational chemistry different?

- Community doesn't care unless they can afford hundreds of jobs *per chemist*.
- We cannot discriminate between the **thousands** of important questions which can be answered by computation.
- The impact of hero simulations is often psychological (moving the flag).

Why is exascale different for chemists?

Exascale is different for chemists because of what happens at petascale.

Petascale is the crossover point between algorithmic approaches for quantum chemistry many-body methods wherein the scalable sparse algorithms overtake the canonical dense algorithms.

Quantum nearsightedness doesn't help if the world is only three inches around!

Even without nearsightedness, massive parallelism opens doors into totally new application areas.

Electronic excited-states in biology

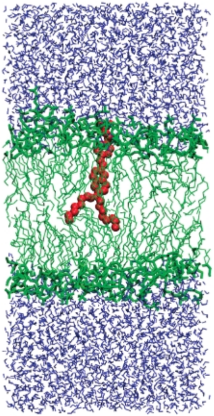
Joint work with Benoît Roux (UC/ANL) and Karol Kowalski (PNNL).

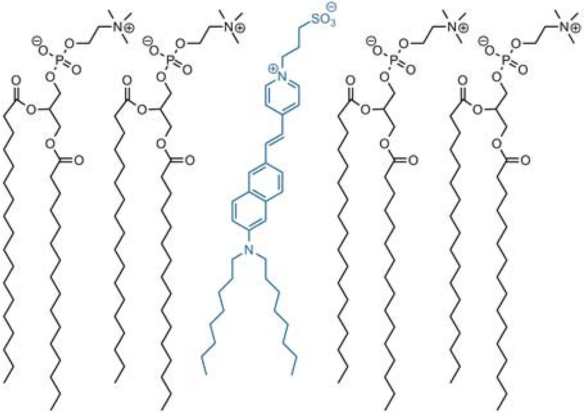
Molecular probes

Optical potentiometric probes have become important tools in electrophysiology. These organic molecules display spectroscopic responses to membrane potential and have been used for the study and characterization of model membranes, nerve and muscle tissues, organelles, microorganisms, and red blood cells. They can often be used in place of conventional microelectrodes and lend themselves to many system not accessible to microelectrodes.

E. Fluhler, V. G. Burnham, L. M. Loew, *Biochemistry* **24** , 5749 (1985). “Spectra, membrane binding, and potentiometric responses of new charge shift probes.”

Membrane configuration of di-8-ANEPPS

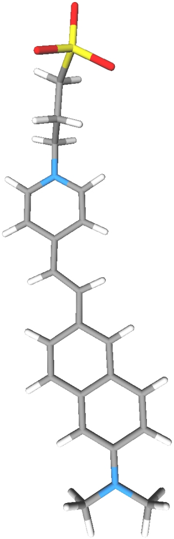


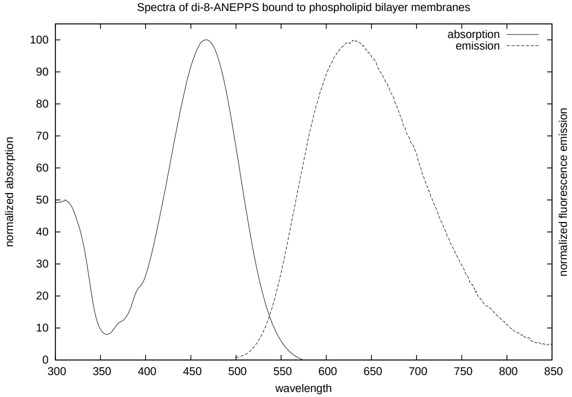


C. F. Rusu, H. Lanig, O. G. Othersen, C. Kryschi, and T. Clark, *J. Phys. Chem. B* **112**, 2445 (2008).

ANEPPS model structure

peaks = 2.655 eV (3.987 eV) and 1.965 eV





`http://probes.invitrogen.com/media/spectra/data/3167lip.txt`

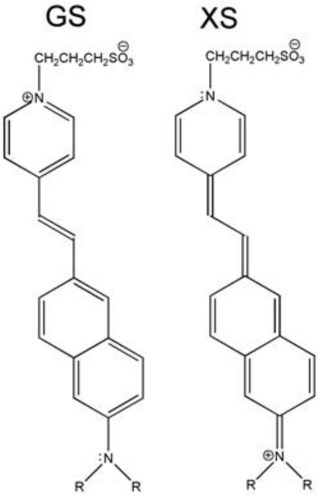
Computing the spectrum

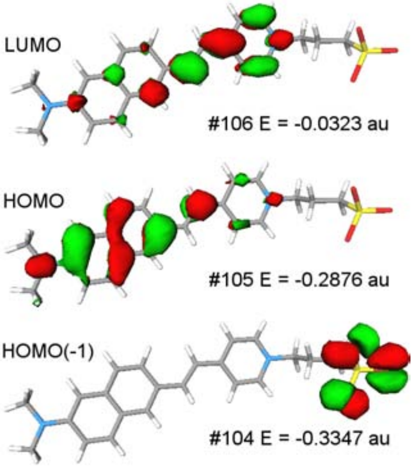
Method comparison

Root	au	eV	nm	OS
B3LYP/cc-pVDZ	0.002	0.06	19758.1	0.001
B3LYP/aug-cc-pVDZ	0.016	0.44	2799.2	0.000
PBE0/aug-cc-pVDZ	0.026	0.70	1773.5	0.000
BH&H/aug-cc-pVDZ	0.091	2.47	501.7	0.000
TDHF/aug-cc-pVDZ	0.124	3.38	366.7	1.731
CIS/aug-cc-pVDZ	0.132	3.59	345.8	1.949

The basis set dependence is an illusion.

What are the electrons doing?





Computing the spectrum

Many-body methods

Basis	τ	eV	au	nm
CC2/6-31G*	1.5	1.815	0.067	683.3
CCSD/6-31G*	0.5	3.629	0.133	341.6
	0.7	3.231	0.119	383.8
	1.5	2.984	0.110	415.5
	2.0	2.962	0.109	418.6
	∞	2.968	0.109	417.7
CCSD/cc-pVDZ	∞	2.945	0.108	421.0
Experiment		2.655	0.098	467

Accurate many-body methods

CR-EOM-CCSD(T) excitation energies

Method	Basis	τ	eV	au	nm
EOM-CCSD	6-31G*	0.5	3.629	0.133	341.6
	6-31G*	0.7	3.231	0.119	383.8
	6-31G*	1.5	2.980	0.110	416.1
CR-EOM-CCSD(T)	6-31G*	0.5	3.590	0.132	345.4
	6-31G*	0.7	3.150	0.116	393.6
	6-31G*	1.5	2.810	0.103	441.2
Experiment			2.655	0.098	467

NWChem implementation (TCE)

Timings on 256 nodes of Chinook

Procedure	wall time (s)
SCF total time	57
four-index transformation	192
one CCSD iteration	157
one EOM-CCSD iteration	252
CR-EOM-CCSD(T) evaluation	6301
Total time	12510

Even though this calculation is trivial with NWChem, it is still impossible with single-node codes because of the memory wall.

Karol is up to 25K cores at NERSC. . .

Bottom-up simulation in material science

Joint work with Karol Kowalski (PNNL).

Graphitic materials

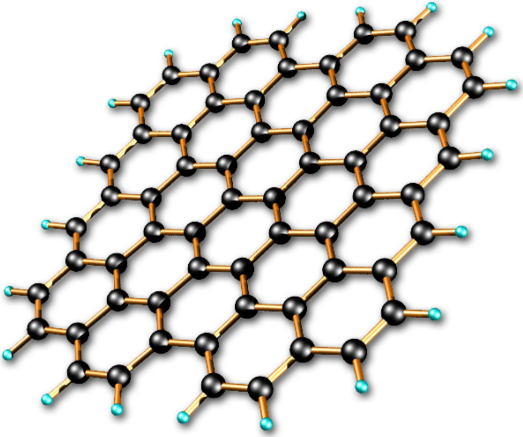
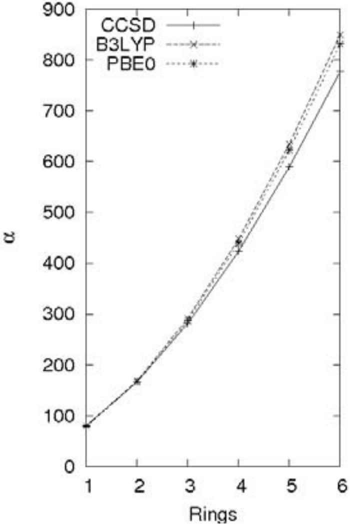


Image from Berkeley Labs (Lanzara Group).

Polarizability simultaneously probes excited-state behavior (poles) and intermolecular forces — dispersion closely related to $\alpha(\omega)$.

Polarizabilities of polyacenes



In June 2006, benzene (1 ring) was the largest CCSD-LR $\alpha(\omega)$ calculation ever performed.

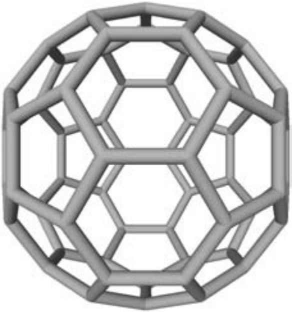
	α_{LL}		
Rings	CCSD	B3LYP	PBE0
1	80.57	79.38	78.75
2	166.61	168.59	166.48
3	281.60	291.56	287.07
4	423.83	447.60	439.52
5	589.97	634.65	622.40
6	776.83	849.55	831.79

J. Chem. Phys. **127**, 144105 (2007).

Polarizabilities of C_{60}

Theory versus experiment

	Wavelength (nm)	
Method	∞	1064
Lowest found	441.3	-
B3LYP/6-31G*	469.0	-
HF/6-31++G	506.8	515.6
Experiment	516.3	533.1
CCSD/Z3Pol	555.3	564.9
LDA/TZP++	571.6	-
CC2/6-31++G	586.8	600.8
CC2/6-31++G*	606.8	622.6
CC2/aug-cc-pVDZ	623.7	640.2
Highest found	1033.2	-



J. Chem. Phys. **129**, 226101 (2008).

Moving the flag

Applying CCSD to C_{60}...

...was laughable in 2000.

...was impossible in 2005.

...was heroic in 2007.

...is mundane in 2010.

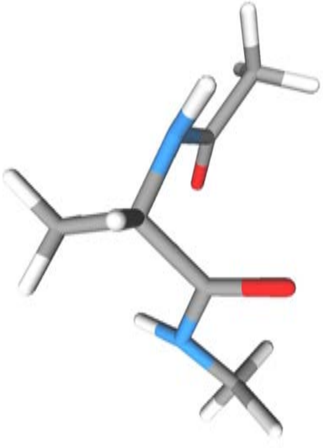
What happened?

- automatic code generation was critical in implementing CCSD-LR in parallel
- finally had a machine that could hold everything in memory
- enough resources such that an intern wasn't afraid to burn millions of hours a year

Force-fields from first-principles

Joint work with Karl Freed (UC), Benoît Roux (UC/ANL), Alex MacKerell (Maryland)

The protein prototype — dialanine



- Debatable if dialanine represents the real torsional potential.
- Many FF potentials use MP2 dialanine results.
- Useful for calibrating methods without pollution of cooperative effects.
- Computationally tractable for CCSD(T) (whole ϕ-ψ map).

Evaluating models with CCSD(T)

Method	MUE	Max
MP2	0.436	1.304
CCSD	0.577	1.426
B2PLYP	0.913	4.690
M06	1.137	4.874
Becke97+D	1.177	5.981
Becke98	1.287	7.526
TPSS	1.312	10.691
B3LYP+D	1.327	6.379
TPSSh	1.330	9.525
M06-L	1.378	6.657
Becke97	1.391	7.486
PBE+D	1.404	9.812
X3LYP	1.430	7.747
B3LYP	1.456	7.884
PBE0	1.506	8.041
M05	1.551	8.389
TPSS+D	1.593	9.258
FT97	1.711	11.047
CAMB3LYP	1.747	6.268
M06-2X	1.757	5.812
BB1K	1.773	7.310
B2LYP	1.913	6.514
HCTH120	2.119	10.141
BOP	2.614	9.118
M06-HF	2.884	12.286
SCF	3.066	11.076
HCTH407	3.168	9.678
HCTH	3.330	9.788
CAMPBE0	3.348	10.676

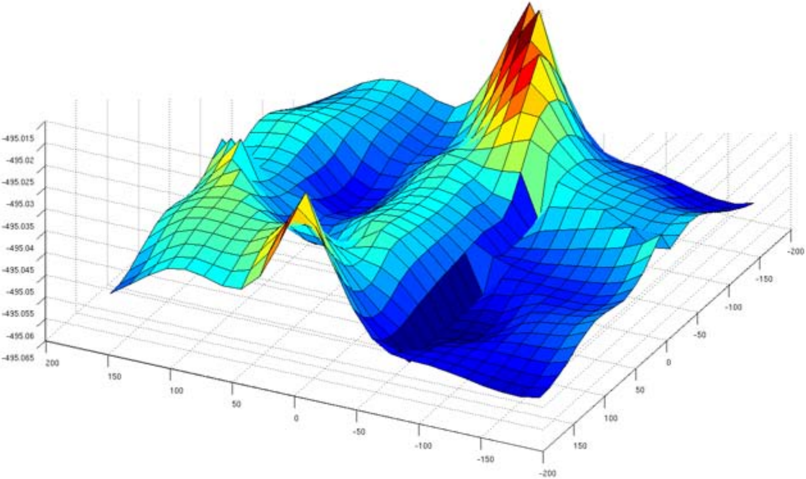
Evaluating models with CCSD(T)

Observations

- Justified using MP2 for fitting torsional parameters.
- Approximate functionals are getting better with time.
- DFT+D improves results in most cases.
- Unlikely that a density functional better than MP2 exists.
- CCSD(T) takes approximately 1 hour per job on 64 nodes.

Endgame for dialanine

CCSD(T)/cc-pVTZ energies at MP2/cc-pVTZ geometries.



Beyond petroleum for the chemical industry

Joint work with Larry Curtiss (ANL) and Jeff Greeley (ANL).

Chemistry after oil

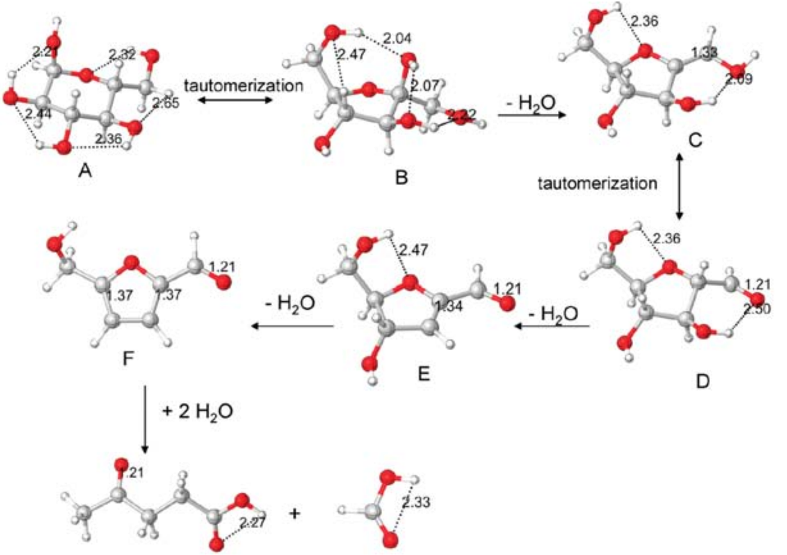
Oil won't disappear, but the price is going to go way up.

Cannot live without plastic and cannot pay more for commodities.

Levulinic acid is a precursor for polymers, plastics and pharmaceuticals.

If we can convert cheap, abundant, non-petroleum chemicals into levulinic acid, there is potential for a chemical industry after oil.

From glucose to levulinic acid



Computational details

The CCSD(T) calculations in the G4 method were prohibitively slow (weeks) using Gaussian, but ran in less than one hour on 1024 nodes of Blue Gene/P using NWChem.

We optimized NWChem CCSD(T) code for Blue Gene/P by developing the first threaded kernels and improving ARMCI.

Larry's G3/G4 methods are the standard model for thermochemistry. If they run on supercomputers rather than workstations, the possible applications grow exponentially.

Summary

Supercomputers and parallel software were critical to the accurate study of four systems:

1. di-8-ANEPPS ion channel probe
2. C_{60} and graphic materials
3. dialanine protein model
4. glucose to levulinic acid

Exascale means the democratization of such capability as well as a paradigm shift in quantum many-body algorithms.

The **Chemistry Exascale Codesign Center** will deliver the transformative software capability required to realize the potential of accurate simulations in many critical areas (biology, material science, energy science).

Extras

Density-functional theory

$$E = T[\rho] + U[\rho] + V_{XC}[\rho]$$

which are the kinetic, Coulomb and exchange-correlation components. In practice, $\rho \leftarrow \{\phi_i\}$ so $T[\rho] = T[\{\phi_i\}]$.

Category	Functional form	Example
LDA	$V_{XC}[\rho_\sigma]$	LDA
GGA	$V_{XC}[\rho_\sigma, \nabla\rho_\sigma]$	BLYP, PBE
Hybrid GGA	$\alpha V_{XC}[\rho_\sigma, \nabla\rho_\sigma] + \beta V_X[\{\phi_i\}]$	B3LYP, PBE0
Meta GGA	$V_{XC}[\rho_\sigma, \nabla\rho_\sigma, \tau_\sigma]$	TPSS
Double-hybrid	$V_{XC}[\rho_\sigma, \nabla\rho_\sigma, \overrightarrow{\epsilon}_\sigma]$	B2PLYP

The exact functional form is not known. The coefficients are usually fit to data but occasionally determined from first-principles.

Coupled-cluster theory

$$
\begin{aligned}
|CC\rangle &= \exp(T)|0\rangle \\
T &= T_1 + T_2 + \cdots + T_n \quad (n \ll N) \\
T_1 &= \sum_{ov} t_o^v \hat{a}_v^\dagger \hat{a}_o
\end{aligned}
$$

- Can do excited-states and arbitrary-order properties
- Fast convergence in T — singles and doubles (CCSD) are an excellent approximation for many problems
- Perturbative corrections, namely CCSD(T), produce extremely accurate results at $n_{iter}N^6 + N^7$ cost ($n_{iter} \approx 20$).
- Memory-bound but highly parallelizable.

Evaluating models with CCSD(T)

- Used OPLS-AA geometries to prevent bias.
- 6-311++G** basis set (aug-cc-pVTZ desirable).
- Difference between 6-31+G* and 6-311++G**: SCF=1.18, MP2=2.48 (MUE in kJ/mol).
- Difference between 6-311++G** and aug-cc-pVTZ: SCF=0.89, MP2=1.70 (MUE in kJ/mol).
- 350 configurations (30° grid everywhere, 10° in basins).